



# User Experience and Perceived Usability of Traditional and Multimodal Interfaces for Scientific Information Seeking

Patrik Pluchino<sup>1</sup>(✉), Luciano Gamberini<sup>1</sup>, and Giulio Jacucci<sup>2</sup>

<sup>1</sup> Human Inspired Technology (HIT) Research Centre,  
University of Padova, Padua, Italy  
patrik.pluchino@unipd.it

<sup>2</sup> Helsinki Institute for Information Technology (HIIT),  
Department of Computer Science, University of Helsinki, Helsinki, Finland

**Abstract.** The current study was carried out to investigate users' preferences for different input modalities (i.e., mouse or gaze) when involved in scientific information-seeking tasks. According to the task that they had to perform (i.e., abstract opening, gaze bookmarking, or keyword highlighting), participants were randomly assigned to three different conditions. They had to carry out the same task twice with both means of interaction (i.e., mouse or gaze); input modalities were counterbalanced across participants. The findings showed that users were faster and perceived the system as more pleasant when they performed the tasks using their gaze. However, when participants controlled the system with a more familiar mouse interface, they perceived the system as more accurate, easy to use, and efficient. Future studies will evaluate how increasing users' familiarity with alternative input modalities (i.e., gaze) can affect their perception of accuracy, easiness, and efficiency.

**Keywords:** Multimodal interface · Eye tracking · Facial expressions · Emotions · Arousal

## 1 Introduction

The term Multimodal Human-Computer Interaction (MMHCI) refers to an emerging field of research that bonds together various domains, such as cognitive psychology, artificial intelligence, computer vision, and many others (Jaimes and Sebe 2007; Turk 2014). The general aim is to create computers that are more usable and that can positively enhance the user experience. Particular consideration has been given to the user, the technological device, and their interaction. In traditional human-computer interaction, users control a system/application using conventional input modalities such as a mouse and/or keyboard. In the context of MMHCI, instead, several different means of control are considered: gestures (Paylovic et al. 1997), voice (Van der Kamp and Sundstedt 2011), eye tracking, and related indices (Grauman et al. 2003; Penkar et al. 2012). An additional crucial point regarding MMHCI is that, similarly to human-human interaction, communication involves a combination of different means (e.g., voice, body and hand gestures, gaze, etc.). Perceptual interfaces (PIs; Turk and

Robertson 2000) are specific systems that aim at providing interaction exploiting non-conventional human-computer interaction input modalities. The latter are more natural and similar to the ones utilized in the real world. Thus, computers that are capable of exploiting these channels of communication can support users in performing various tasks. These machines can implement actions on the basis of users' gaze position on the screen, or they can modify their functioning in accordance with the emotions that users' faces show. Jaimes and Sebe (2007) characterized MMHCI in three categories according to the different parts of the human body involved: body movements, hand gestures, and gaze. The first category is related to technological systems that aim at interpreting general body posture and motion in order to change their operating behavior. The second category refers to devices that are able to adapt their functioning based on codified essential actions (e.g., pointing at specific objects) or on more complex ones (e.g., actions that convey feelings). The third category concerns the exploitation of eye-related indices and is based on the eye-mind hypothesis (Just and Carpenter 1976), which states that the spatial position of the gaze reflects what the participant is currently processing. This latter category is the more suitable in the context of utilization and evaluation of desktop interfaces. One key feature of the eye-tracking methodology is that it permits tracing of cognitive processing in an unobtrusive way. Different types of data (e.g., fixation and saccade durations, dwell time, etc.) can be collected while participants are dealing with different kinds of stimuli (e.g., texts, images, visual search, scene perception; Rayner 1998, 2009). Nevertheless, the eye tracker can also be exploited as a means of control. Indeed, it is possible to acquire eye-related data in real time, and computers can utilize this information to implement simple actions (e.g., selecting a button on the screen). Duchowski (2017) described such tools as selective interactive systems when changes that occur on the display are the consequence of an intentional gaze pattern. Indeed, in these conditions, the gaze is utilized to operate the systems instead of traditional controllers such as a mouse, a joystick, and/or a keyboard. Several studies have focused on disabled users (Levine 1984; Bates et al. 2007; Donegan et al. 2009), but eye tracking can also be adopted as a means of interaction for non-disabled people (Hyrskykari et al. 2005; Drewes 2010). Moreover, it is important to bear in mind that there could be constraints, for instance, that the visual modality in such conditions will be utilized for both perceiving the stimuli and controlling the interface/system. For this reason, the implementation of such a means of control has to be capable of discriminating between situations of "casual looking" and intentional viewing. The purpose is to prevent the "Midas touch" issue that concerns the selection of items that are randomly fixated on and not just the ones relevant to the current task that has to be accomplished (Jacob 1991; Majaranta and Bulling 2014).

The current experiment aimed at comparing, in terms of performance and user experience, the gaze input modality with a more conventional type of control, namely the mouse, while users performed different information-seeking tasks. Moreover, users' experienced emotions and level of general activation (arousal) were monitored during the interaction.

2 Method

2.1 Participants

Fifty-nine students from the University of Padua took part in the experiment on a voluntary basis. Eight participants were excluded as outliers due to the following: long temporal session duration, poor quality of the video recordings, and the impossibility of using the recordings to analyze participants’ emotions and arousal. Fifty-one students ( $F = 23$ ) with a mean age of 27.05 years ( $SD = 3.87$ ) were considered for the analyses. All participants presented normal or corrected-to-normal vision. Seventeen participants were randomly assigned to each of the three experimental conditions (see *Experimental conditions and tasks* section for details).

2.2 Materials

**Stimuli.** The interface was developed in C++. Each screen presented a query field at the top left; an *orbitarium* at the center, where scientific keywords were graphically presented using polar coordinates; and finally, a set of abstract papers on the right side (see Fig. 1). The interface is described in detail by Serim and collaborators (Serim et al. 2017). When a/more word/s was/were inserted in the query field, a search iteration could be started by pressing the enter key. Thus, a pool of abstracts was selected from the interface database. This database retrieved over 50 million scientific documents from the Web of Science prepared by Thomson Reuters, Inc., and from the Digital Libraries of the Association of Computing Machinery (ACM), the Institute of Electrical and Electronics Engineers (IEEE), and Springer. Simultaneously, from among all the keywords that belonged to these selected abstracts, the more relevant keywords were shown inside the *orbitarium*. Furthermore, the relative positions of the keywords determined their relevance (i.e., more relevant keywords were located closer to the center).

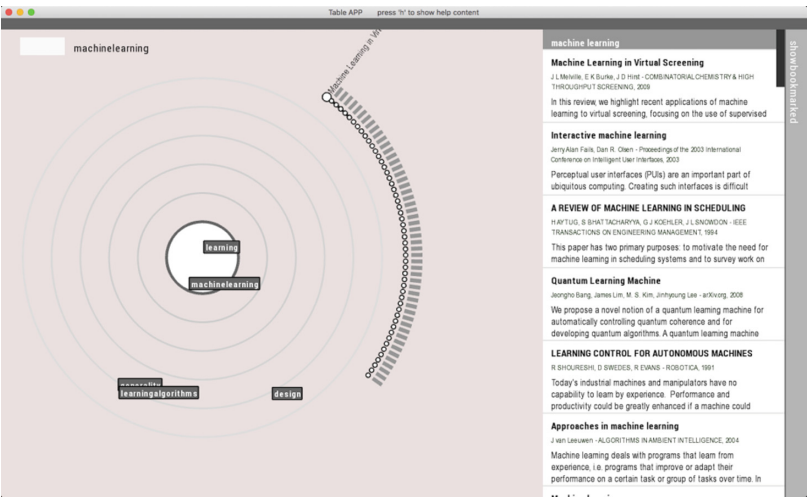


Fig. 1. The interface for scientific information seeking.

**Experimental Conditions and Tasks.** Regardless of the experimental condition, participants performed the same task twice utilizing the mouse and their gaze as means of control. The order was counterbalanced across participants. In each condition, the experimenter entered a topic keyword in the query field that started the first search iteration. On the basis of the resulting screen, participants had to, respectively, open 7 abstracts (*abstract opening* condition); bookmark 7 abstracts (*gaze bookmarking* condition); or access all the keywords that were linked to 7 abstracts (*fade in and out* condition). In all conditions, participants had to choose a set of abstracts that were relevant to the initial topic keyword. Participants could scroll up/down the list of abstracts in both experimental sessions (mouse and gaze-controlled).

**Equipment.** Several devices were adopted in the present experiment (Fig. 2). A RED500 (SMI) remote eye tracker with a sampling frequency of 500 Hz was utilized to collect the eye-tracking data. The software iViewX v. 2.8 (SMI), which was installed on a Dell Latitude E6530 notebook, was utilized to store the eye-tracking data. The software Experiment Center v. 3.6 (SMI), which was installed on a Dell desktop computer, was utilized to perform and validate the calibration of the eye tracker. The interface for scientific information seeking was installed on a MacPro that was connected to a 22" Dell monitor (1680 × 1050 pixels) under which the eye tracker was placed. Participants' faces were video recorded during the experimental sessions. The quality of the videos was enhanced utilizing two halogen bulb lights (60 W) that were located at both ends of the screen. Finally, utilizing the FaceReader software (Noldus), the recorded videos were analyzed to compute the proportion of different participants' emotions and arousal levels. The software is capable of automatically analyzing six basic facial expressions (i.e., happiness, sadness, fear, disgust, surprise, and anger) and a neutral state. Furthermore, FaceReader provides the opportunity to measure the emotional valence (pleasant = 1, unpleasant = -1) and the arousal level (active = 1, inactive = 0). All the computers were connected through a TL-SF1005D Ethernet switch. Thus, data regarding the gaze positions, collected by iView X, were sent to the C++ interface to enable the participants to accomplish the tasks using their eyes.

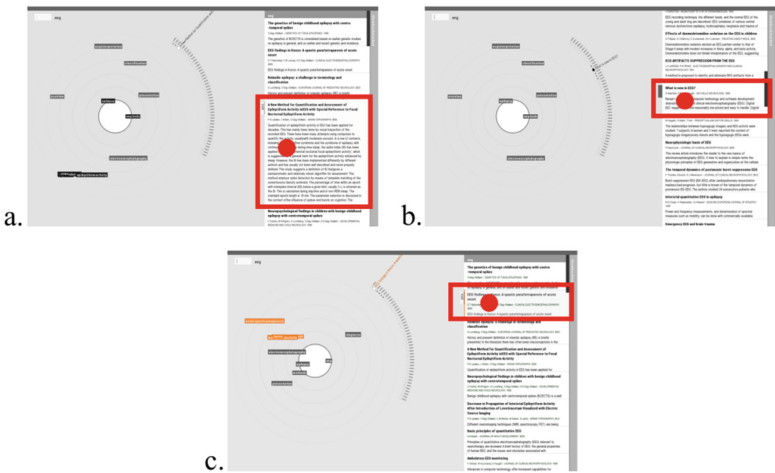
**Experimental Design.** The independent manipulation of two factors was considered (mixed-design): The *within*-participants factor was the means of control (gaze vs. mouse), while the *between*-participants factor was the kind of task that participants had to accomplish (abstract opening vs. gaze bookmarking vs. keyword fade in and out).

**Procedure.** Upon arrival, participants were welcomed and asked to read and potentially sign an informed consent form. In the experimental setting, participants were seated around 60–70 cm away from the screen. They were instructed to find a comfortable position that they would have to maintain during the experiment. Indeed, they had to try to avoid, as much as possible, head and body movements. In the *mouse control* experimental session, participants interacted with the interface by utilizing the mouse; the topic keyword was *database*. In the *gaze control* experimental session, users controlled the interface using their gaze; the topic keyword was *machine learning*. Before the *gaze control* session, the eye tracker was calibrated (accuracy: 0.5° visual angle) using a 5-point calibration procedure. With this means of control, according to the specific task, participants had to fixate on an abstract for at least three



**Fig. 2.** Experimental setup.

seconds (non-cumulative) in order to open it (*abstract opening* condition, Fig. 3a); fixate on an abstract for at least 5 s (non-cumulative) in order to bookmark it (*gaze bookmarking* condition, Fig. 3b); or fixate on an abstract for a non-specified amount of time to make the correspondent keywords appear (*fade in and out* condition, Fig. 3c). A training session preceded both experimental sessions. The aim was to give participants the opportunity to become familiar with the interface and its controls. In the training phases, the topic keywords were *EEG* and *ERPs*. At the end of both



**Fig. 3.** Experimental tasks: (a) *abstract opening*, (b) *gaze bookmarking*, and (c) *fade in and out* utilizing the gaze (the red dot corresponds to the gaze position; it is shown for clarity and was not presented to the participants). (Color figure online)

experimental sessions, participants were administered an electronic questionnaire regarding the interface evaluation. Then, participants were debriefed about the study aims. Both experimental sessions together lasted 25 min.

**Measures.** Several metrics were collected during and after the experimental sessions:

- *Experimental session duration.* The total time (sec) needed to perform a task.
- *Experienced emotions.* The percentage of positive and negative emotions experienced during each experimental session.
- *Emotional valence.* The maximal level of emotional valence (range: -1 to 1) reached during each experimental session.
- *Arousal.* The maximal level of arousal (range: 0 to 1) reached during each experimental session.
- *System evaluation.* An ad hoc 20-item user experience questionnaire was administered (10 items of evaluation regarding each means of control). Usability (easiness, efficiency, fatigue, clarity, speed, fluidity, intuitiveness), pleasantness, perceived utility, and accuracy of the interface were evaluated using a 5-point scale (1 = not at all; 5 = very; see Appendix for detailed information about the items).

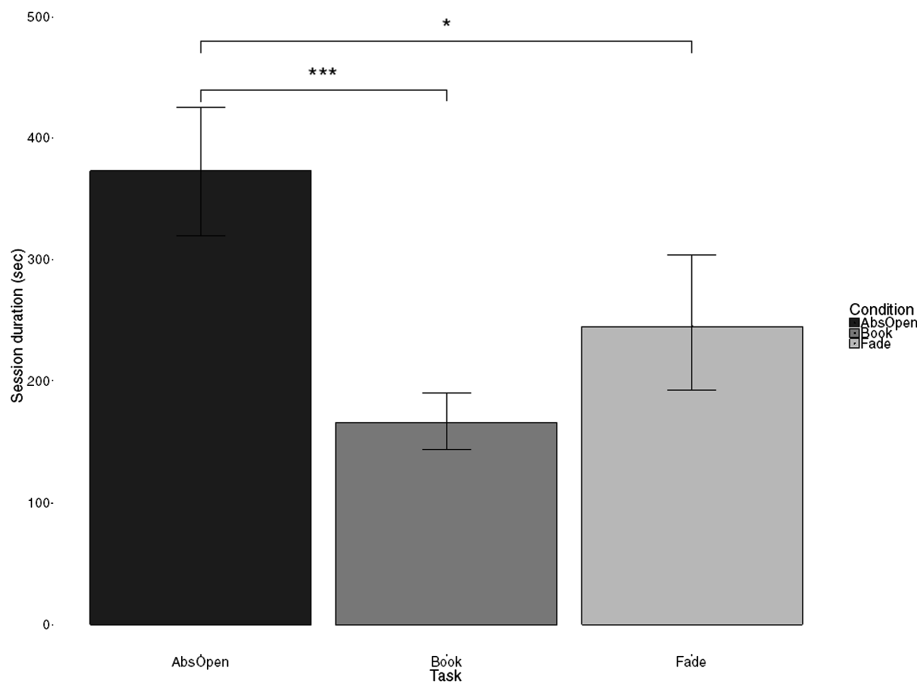
### 3 Data Analyses

The interquartile range (IQR) procedure (i.e., all values falling outside 1.5 IQR from the extremes of the IQR box were considered outliers) was adopted to exclude participants that were outliers in terms of total duration in at least one of the two experimental sessions. Some of the following analyses were conducted by means of mixed models (generalized in accordance with non-normal distribution of the data). The considered fixed effects were the task (abstract opening vs. gaze bookmarking vs. keyword fade in and out) and the type of control (mouse vs. gaze) utilized to accomplish the assigned task. Participants were considered a random effect. Abstract opening and mouse were set as the contrast levels. These analyses were performed using the R package lme4 (R Core Team 2015). FaceReader data were pre-processed utilizing a set of customized MATLAB functions (Release 2015a, Mathworks Inc.). Following the pre-processing, these data were analyzed through two beta regression analyses (positive and negative percentage of experienced emotions) utilizing the R package betareg (Cribari-Neto and Zeileis 2010).

#### 3.1 Results

**Experimental Session Duration (Generalized Mixed-Models).** A main effect of task type emerged. Users were faster in accomplishing the abstract bookmarking task ( $b = -0.82$ ,  $t = -4.05$ ,  $p < .001$ ;  $M = 165.76$  s) when they had to access all the keywords of the selected abstracts ( $b = -0.49$ ,  $t = -2.41$ ,  $p < .05$ ;  $M = 244.67$  s) compared to when they had to open the abstracts ( $M = 373.07$  s; see Fig. 4).

Moreover, a main effect of the means of control was found ( $b = -0.15$ ,  $t = -2.15$ ,  $p < .05$ ). Users in general needed a shorter amount of time to perform the various tasks



**Fig. 4.** Mean duration of an experimental session as a function of task.

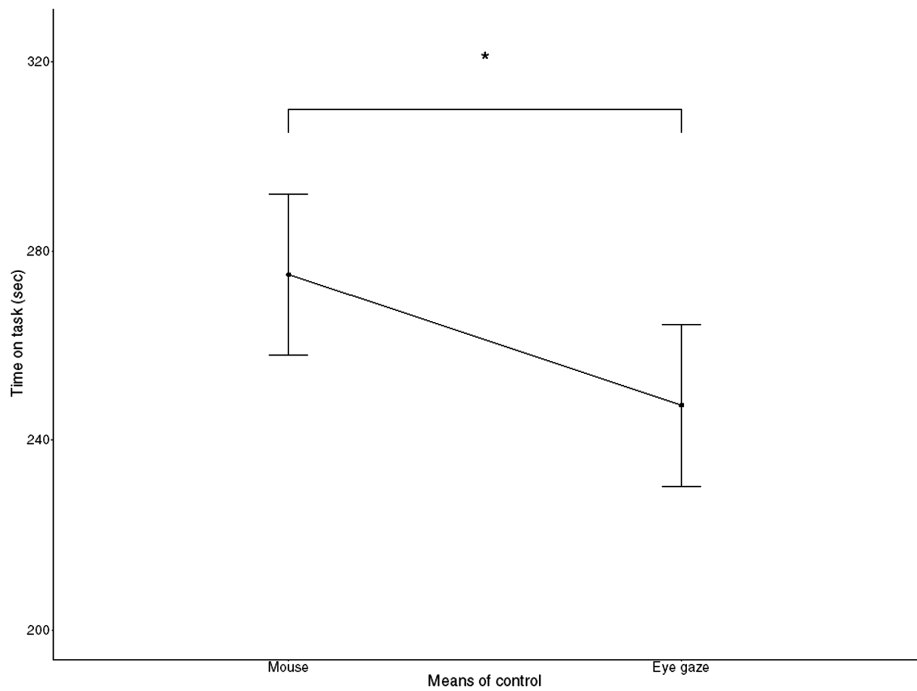
utilizing their gaze ( $M = 247.34$  s) compared to when they interacted with the system using the mouse ( $M = 274.99$  s; Fig. 5). No significant interaction emerged.

**Experienced Emotions (Beta Regressions).** No main effect emerged considering either the positive or the negative emotions. Similar percentages were shown despite both the main effect of task type and of means of control (see Table 1 for positive emotions and Table 2 for negative emotions).

**Emotional Valence.** A main effect of the means of control on the maximal level of emotional valence emerged ( $b = 0.07$ ,  $t = 2.80$ ,  $p < .05$ ; see Fig. 6). This value was higher when participants were controlling the interface with their gaze ( $M = 0.83$ ) compared to when they utilized the mouse ( $M = 0.73$ ). No significant main effect emerged considering the main effect of task type.

**Arousal.** No main effect emerged considering the maximal arousal level insofar as it was similar despite the main effect of both the task type and the means of control (see Table 3).

**System Evaluation.** A series of Kruskal-Wallis tests were performed considering delta values to evaluate the main effect of the task. Delta values were computed for each pair of scores regarding the same item (e.g., pleasantness: mouse score – gaze score). No differences emerged.



**Fig. 5.** Average duration of an experimental session as a function of means of control.

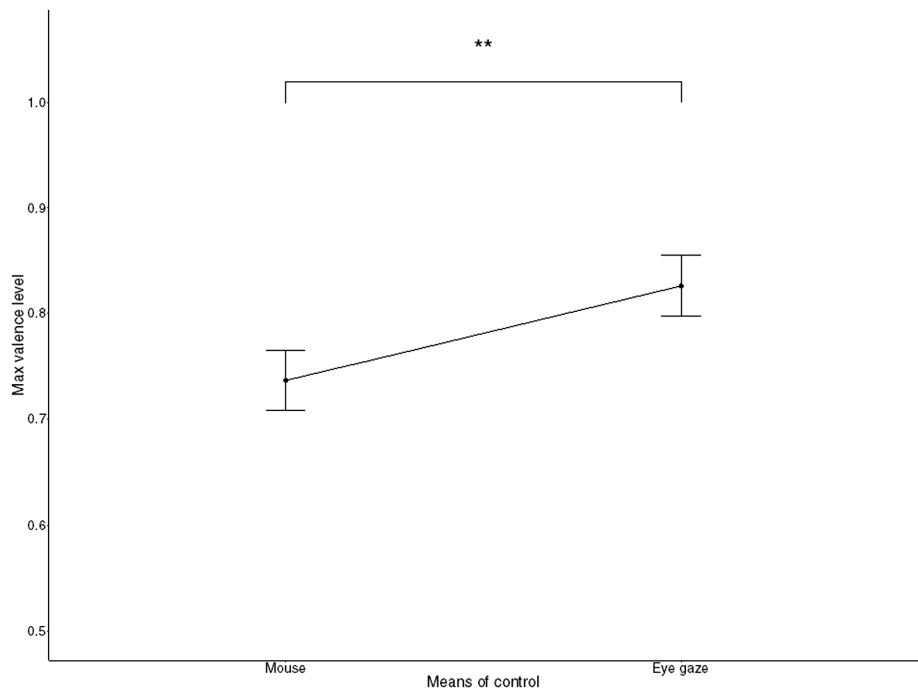
**Table 1.** Percentage of positive emotions as a function of task and means of control.

Tasks	Controls	
	Mouse	Eye gaze
	<i>M(SD)</i>	<i>M(SD)</i>
Abs opening	13.59(17.40)	14.40(18.33)
Bookmarking	9.52(10.93)	12.38(15.78)
Fade in/out	17.61(18.63)	14.68(17.32)

**Table 2.** Percentage of negative emotions as a function of task and means of control.

Tasks	Controls	
	Mouse	Eye gaze
	<i>M(SD)</i>	<i>M(SD)</i>
Abs opening	42.64(24.09)	37.95(22.16)
Bookmarking	53.54(32.93)	47.12(34.81)
Fade in/out	30.70(24.56)	36.43(29.05)



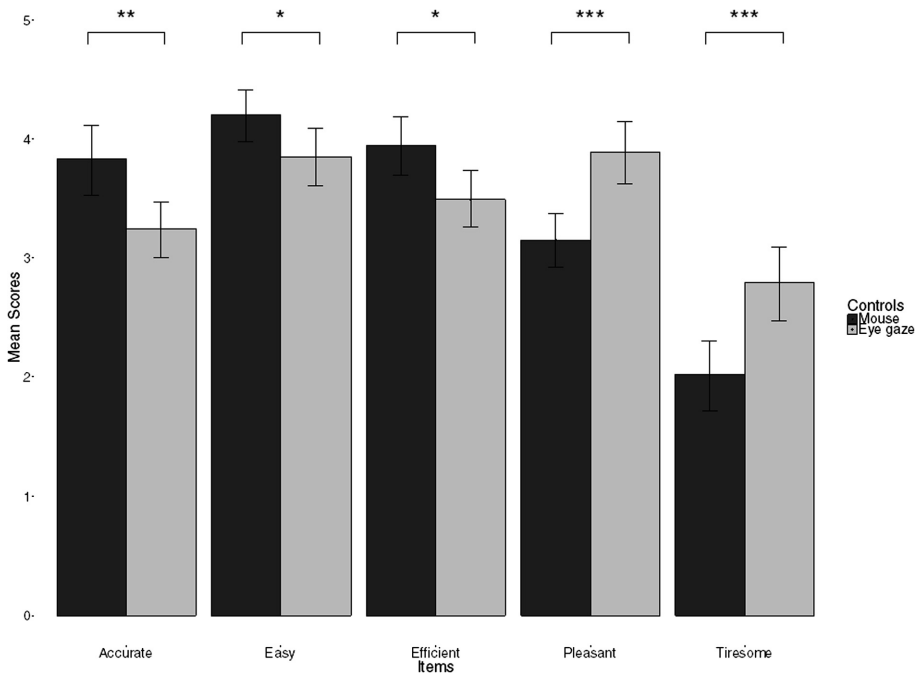


**Fig. 6.** Maximal emotional valence level as a function of means of control.

**Table 3.** Maximal arousal level as a function of task and means of control.

Tasks	Controls	
	Mouse	Eye gaze
	<i>M(SD)</i>	<i>M(SD)</i>
Abs opening	0.66(0.13)	0.60(0.13)
Bookmarking	0.59(0.13)	0.56(0.16)
Fade in/out	0.59(0.12)	0.58(0.13)

Furthermore, a series of Wilcoxon tests were conducted to evaluate the main effect of the means of control. Some differences in the questionnaire scores emerged (Fig. 7). Users perceived the interface as more accurate ( $W = 1743$ ,  $p < .01$ ), more efficient ( $W = 1697$ ,  $p < .05$ ), easier ( $W = 1626.5$ ,  $p < .05$ ), and less tiresome ( $W = 868.5$ ,  $p < .001$ ) when they interacted with it utilizing the mouse. In contrast, participants experienced a more pleasant interaction ( $W = 725$ ,  $p < .001$ ) when the means of control was the gaze.



**Fig. 7.** Mean scores for the questionnaire items that showed differences as a function of means of control.

## 4 General Discussion

The aim of the study was to investigate how the input modality (mouse vs. gaze) could influence the performance and user experience of participants, as well as their emotions, when they were performing information-seeking tasks.

The results in terms of performance (i.e., time needed to accomplish a task) showed that participants were faster in carrying out the various tasks when they were utilizing their gaze as a means of control. These outcomes are in agreement with the idea that eye-related input is quicker than other forms of input, such as mechanical pointing devices (Jacob and Karn 2003). Another study showed that the action of pointing using the eyes was almost 10 times faster than utilizing a joystick in an immersive projection display (Asai et al. 2000). The authors stated that an eye-pointing task was not influenced by participants' experience when it was performed using the eyes. Indeed, individuals could easily and similarly (in terms of time on task) point using their eyes in the absence of specific training. In contrast, the large variability of time needed when utilizing the joystick was related to the participants' level of practice. Moreover, Jacob (1991) found a decrease of 30% in the time needed to accomplish a selection task utilizing the eyes compared to the mouse. Likewise, Tanriverdi and Jacob (2000) showed an advantage of gaze, compared to the hand, in terms of speed in carrying out a selection task. Moreover, any hand-controlled input tool could implement a specific

action (e.g., mouse cursor movement) just on the basis of the previous eye movements made on the destination of this specific goal-oriented action.

Regarding the emotions experienced, no differences emerged for positive or for negative emotions. These results could be expected because the interface proposed had, *per se*, a low emotional valence. The same applies to the outcomes of the analysis concerning the maximal arousal level. To support the accuracy in collecting data on arousal level of participants utilizing the FaceReader software, the utilization of surface electrodes for monitoring the electrodermal activity could be considered. A difference emerged concerning the maximal level of emotional valence. Results showed that participants experienced a higher peak of positive emotions when they were interacting with the interface through their gaze. Considered together with the higher level of perceived pleasantness of the interface, in the gaze control experimental sessions, this occurrence confirmed the idea that the eye tracker could be exploited as a means of interaction and not only as a tool that records eye-related data in a passive way. Indeed, participants considered positively interacting through this alternative means of control at both the quantitative (maximal valence) and subjective level (evaluation in terms of pleasantness).

The findings about the user experience underlined issues related to the fact the participants are not accustomed to interfaces that can be controlled with their eyes (Jacob and Karn 2003). Compared to a traditional input modality (i.e., mouse) that gives constant visual feedback to the users regarding the actual position of the cursor on the screen, gaze control does not provide such information. This could explain why participants perceived the interface as less accurate, efficient, easy to use, and tiresome. Indeed, users did not know where their gaze was precisely located at any time, and this uncertainty led them to subjectively perceive the interface functioning in a negative way, in accordance with previous works (Serim et al. 2017). The interface only provided raw feedback about gaze data availability. Indeed, the color of the interface background slightly changes from the original grey when the eye tracker is not properly collecting gaze data (i.e., overly extensive head/body movements outside of the so-called head box; Serim et al. 2017). No actual point of gaze indicator was continuously shown. The authors made this choice with the aim of keeping the interaction as natural as possible, in line with previous findings (Slobodenyuk 2016). Indeed, the cursor would hamper the task accomplishment insofar as it would partially hide the information at which the participant was looking. Nevertheless, to experience the so-called oculomotor agency, which Slobodenyuk (2016) defined as “an experience of control over eye movements that involves perception of gaze-related causality and correspondence of the outcome to intention,” the presence of a cursor that indicates the actual gaze point is necessary. This is in accordance with a literature study that demonstrated how the self-agency of participants was very low when the cursor was absent (Wang et al. 2012). Participants could not explicitly experience the sense of gaze agency, which occurrence is reflected in the negative perception of the interface. Moreover, eyes are characterized by continuous small movements, even during a fixation (i.e., microsaccades, tremors, and drifts) that can be additional factors of gaze-location instability. Finally, individuals are not familiar with real-world objects that respond to their eye movements. The only exception is when they are interacting with other people (Jacob and Karn 2003).

In general, a pivotal aspect could be clarifying the eye tracker functioning insofar as head and body movements adversely affect its tracking accuracy. Morimoto and Mimica (2005) pointed out how such issues impact the usability and, as a consequence, wider uptake of gaze-based interactive systems. During the *gaze control* experimental sessions, participants might forget this constraint. Thus, the potential decrease in tracking accuracy could cause them to negatively perceive the interface responsiveness. In contrast, the interface was perceived as more pleasant when users performed the task by means of their gaze. The interface could be perceived as capable of understanding participants' intentions before they actually express them (Jacob 1991). Moreover, this could mean that the implicit feature of this technique (the participants do not continuously pay attention to the fact that the eye tracker is working) can positively affect the user experience evaluation.

Considering the overall findings, it is possible to speculate on the opportunity to control interfaces by adopting a MMHCI approach combining, for instance, a traditional modality (i.e., mouse and keyboard) with one or more alternative modalities of interaction (e.g., gaze). Nevertheless, it is crucial to bear in mind that the interfaces or the target applications have to be conceived on the basis of the alternative modalities' characteristics. In a multimodal interface/system that exploits the eye direction as an alternative input modality, the size of the gaze-interactive item (e.g., button or menu) has to be large enough to allow a fluid interaction. This constraint could force, for instance, creation of menus and submenus, which are well-known factors that can reduce the interaction flow. Differently, the size of the mouse-interactive items (e.g., scroll bar) does not need to be changed. This kind of advanced interface has to show a display that contains a combination of stimuli that show *gaze*-friendly features and stimuli that show *mouse*-friendly characteristics. This could ensure full exploitation of the strength points of both input modalities. Moreover, the outcomes regarding the higher perceived pleasantness of the gaze-controlled version of the interface could be exploited. Users should be provided with full information about the issues that can adversely affect the eye-tracking accuracy and should be trained on how to avoid them. An expected consequence will be better evaluation of the gaze-based component of these envisioned multi-modal interfaces in terms of accuracy, efficiency, and ease of use. In these cutting-edge interfaces, the combination of different input modalities will lead to more effective human-computer communication that will resemble human-human communication (Jaimes and Sebe 2007).

**Acknowledgements.** This work was supported by the European Commission (Symbiotic Mind Computer Interaction for Information Seeking, MindSee FP7-ICT; Grant Agreement #611570).

## Appendix

### *User Experience Questionnaire*

Considering the situation in which you control the interface utilizing mouse/gaze control, please indicate using a 5-point scale (1 = not at all; 5 = very):

1. ...to what extent you believe the interface to be easy
2. ...to what extent you believe the interface to be intuitive
3. ...to what extent you believe the interface to be pleasant
4. ...to what extent you believe the interface to be useful
5. ...to what extent you believe the interface to be tiresome
6. ...to what extent you believe the interface to be fluid
7. ...to what extent you believe the interface to be accurate
8. ...to what extent you believe the interface to be efficient
9. ...to what extent you believe the interface to be clear
10. ...to what extent you believe the interface to be fast.

## References

- Asai, K., et al.: Eye mark pointer in immersive projection display. *Proc. IEEE Virtual Reality* **2000**, 125–132 (2000)
- Bates, R., Donegan, M., Istance, H.O., Hansen, J.P., Rähä, K.J.: Introducing COGAIN: communication by gaze interaction. *Univ. Access Inf. Soc.* **6**(2), 159–166 (2007)
- Cribari-Neto, F., Zeileis, A.: Beta regression in R. *J. Stat. Softw.* **34**(2), 1–24 (2010)
- Donegan, M., et al.: Understanding users and their needs. *Univ. Access Inf. Soc.* **8**(4), 259 (2009)
- Drewes, H.: Eye gaze tracking for human computer interaction (Doctoral dissertation, LMU) (2010)
- Duchowski, A.T.: *Eye Tracking Methodology: Theory and Practice*, p. 366. Springer, Heidelberg (2017). <https://doi.org/10.1007/978-3-319-57883-5>
- Grauman, K., Betke, M., Lombardi, J., Gips, J., Bradski, G.R.: Communication via eye blinks and eyebrow raises: video-based human-computer interfaces. *Univ. Access Inf. Soc.* **2**(4), 359–373 (2003)
- Hyrskykari, A., Majaranta, P., Rähä, K.J.: From gaze control to attentive interfaces. In: *Proceedings of HCII*, vol. 2, July 2005
- Jaimes, A., Sebe, N.: Multimodal human–computer interaction: a survey. *Comput. Vis. Image Understand.* **108**(1), 116–134 (2007)
- Jacob, R.J.: The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Trans. Inf. Syst. (TOIS)* **9**(2), 152–169 (1991)
- Jacob, R.J., Karn, K.S.: Eye tracking in human-computer interaction and usability research: ready to deliver the promises. In: Hyönä, J., Radach, R., Deubel, H. (eds.) *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, pp. 573–605. Elsevier Science, Amsterdam (2003)
- Just, M.A., Carpenter, P.A.: Eye fixations and cognitive processes. *Cogn. Psychol.* **8**(4), 441–480 (1976)
- Levine, J.L.: Performance of an eye tracker for office use. *Comput. Biol. Med.* **14**(1), 77–89 (1984)
- Majaranta, P., Bulling, A.: Eye tracking and eye-based human–computer interaction. In: Fairclough, S.H., Gilleade, K. (eds.) *Advances in Physiological Computing*. HIS, pp. 39–65. Springer, London (2014). [https://doi.org/10.1007/978-1-4471-6392-3\\_3](https://doi.org/10.1007/978-1-4471-6392-3_3)
- Morimoto, C.H., Mimica, M.R.: Eye gaze tracking techniques for interactive applications. *Comput. Vis. Image Understand.* **98**(1), 4–24 (2005)
- Pavlovic, V.I., Sharma, R., Huang, T.S.: Visual interpretation of hand gestures for human-computer interaction: a review. *IEEE Trans. Pattern Anal. Mach. Intell.* **7**, 677–695 (1997)

- Penkar, A.M., Lutteroth, C., Weber, G.: Designing for the eye: design parameters for dwell in gaze interaction. In: Proceedings of the 24th Australian Computer-Human Interaction Conference, pp. 479–488. ACM, November 2012
- Rayner, K.: Eye movements in reading and information processing: 20 years of research. *Psychol. Bull.* **124**(3), 372 (1998)
- Rayner, K.: Eye movements and attention in reading, scene perception, and visual search. *Q. J. Exp. Psychol.* **62**(8), 1457–1506 (2009)
- R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2015). <http://www.R-project.org/>
- Serim, B., Chech, L., Vasilieva, M., Papa, L., Gamberini, L., Jacucci, G.: Exploring gaze-adaptive features for interacting with multi-document visualizations. In: Proceedings of the 10th International Symposium on Visual Information Communication and Interaction, pp. 85–92. ACM, August 2017
- Slobodenyuk, N.: Towards cognitively grounded gaze-controlled interfaces. *Pers. Ubiquit. Comput.* **20**(6), 1035–1047 (2016)
- Tanriverdi, V., Jacob, R.J.: Interacting with eye movements in virtual environments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 265–272. ACM, April 2000
- Turk, M., Robertson, G.: Perceptual user interfaces (introduction). *Commun. ACM* **43**(3), 32–34 (2000)
- Turk, M.: Multimodal interaction: a review. *Pattern Recogn. Lett.* **36**, 189–195 (2014)
- Van der Kamp, J., Sundstedt, V.: Gaze and voice controlled drawing. In: Proceedings of the 1st Conference on Novel Gaze-Controlled Applications, p. 9. ACM, May 2011
- Wang, Q., Bolhuis, J., Rothkopf, C.A., Kolling, T., Knopf, M., Triesch, J.: Infants in control: rapid anticipation of action outcomes in a gaze-contingent paradigm. *PLoS ONE* **7**(2), e30884 (2012)